

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Robert G. Ulrich
- (ii) TITLE OF INVENTION: Bacterial Superantigen Vaccines
- (iii) NUMBER OF SEQUENCES: 31
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Elizabeth Arwine
 - (B) STREET: US Army MPMC -504 Scott Street
MCMR-JA (Elizabeth Arwine-Patent
Atty)
 - (C) CITY: FORT DETRICK
 - (D) STATE: MARYLAND
 - (E) COUNTRY: USA
 - (F) ZIP: 21702-5012
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: Apple Macintosh
 - (C) OPERATING SYSTEM: Macintosh 7.5
 - (D) SOFTWARE: Microsoft Word 6.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/882,431
 - (B) FILING DATE: June 25, 1997
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sana A. Pratt
 - (B) REGISTRATION NUMBER: 39,441
 - (C) REFERENCE/DOCKET NUMBER:
- (ix) TELECOMMUNICATION INFORMATION
 - (A) TELEPHONE: (301) 619-2065
 - (B) TELEFAX: (301) 619-7714

(2) INFORMATION FOR SEQUENCE ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 830
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Unknown
 - (D) TOPOLOGY: Unknown

(D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAAAAAAAA CAGCATTAC ATTACTTTTA TTCATTGCCC	40
TAACGTTGAC AACAAAGTCCA CTTGTAAATG GTAGCGAGAA	80
AAGCGAAGAA ATAAATGAAA AAGATTTGCG AAAAAAGTCT	120
GAATTGCAGG GAACAGCTTT AGGCAATCTT AAACAAATCT	160
ATTATTACAA TGAAAAAGCT AAAACTGAAA ATAAAGAGAG	200
TCACGATCAA TTTCGACAGC ATACTATATT GTTTAAAGGC	240
TTTTTTACAG ATCATTCGTG GTATAACGAT TTATTAGTAC	280
GTTTTGATTC AAAGGATATT GTTGATAAAT ATAAAGGGAA	320
AAAAGTAGAC TTGTATGGTG CTTATGCTGG TTATCAATGT	360
GCGGGTGGTA CACCAAACAA AACAGCTTGT ATGTATGGTG	400
GTGTAACGTT ACATGATAAT AATCGATTGA CCGAAGAGAA	440
AAAAGTGCCG ATCAATTTAT GGCTAGACGG TAAACAAAAT	480
ACAGTACCTT TGGAAACGGT TAAAACGAAT AAGAAAAATG	520
TAACTGTTCA GGAGTTGGAT CTTCAAGCAA GACGTTATTT	560
ACAGGAAAAA TATAATTTAT ATAACCTCTGA TGTTTTGTAT	600
GGGAAGGTTT AGAGGGGATT AATCGTGTTT CATACTTCTA	640
CAGAACCTTC GGTTAATTAC GATTTATTTG GTGCTCAAGG	680
ACAGTATTCA AATACACTAT TAAGAATATA TAGAGATAAT	720
AAAACGATTA ACTCTGAAAA CATGCATATT GATATATATT	760
TATATACAAG TTAAACATGG TAGTTTTGAC CAACGTAATG	800
TTCAGATTAT TATGAACCGA GAATAATCTA	830

(3) INFORMATION FOR SEQUENCE ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 257

(B) TYPE: Amino Acid
 (C) STRANDEDNESS: Unknown
 (D) TOPOLOGY: Unknown
 (ii) Molecule type: Peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Lys	Lys	Thr	Ala	Phe	Thr	Leu	Leu	Leu	1	5	10
Phe	Ile	Ala	Leu	Thr	Leu	Thr	Thr	Ser	Pro	15	20	
Leu	Val	Asn	Gly	Ser	Glu	Lys	Ser	Glu	Glu	25	30	
Ile	Asn	Glu	Lys	Asp	Leu	Arg	Lys	Lys	Ser	35	40	
Glu	Leu	Gln	Gly	Thr	Ala	Leu	Gly	Asn	Leu	45	50	
Lys	Gln	Ile	Tyr	Tyr	Tyr	Asn	Glu	Lys	Ala	55	60	
Lys	Thr	Glu	Asn	Lys	Glu	Ser	His	Asp	Gln	65	70	
Phe	Arg	Gln	His	Thr	Ile	Leu	Phe	Lys	Gly	75	80	
Phe	Phe	Thr	Asp	His	Ser	Trp	Tyr	Asn	Asp	85	90	
Leu	Leu	Val	Arg	Phe	Asp	Ser	Lys	Asp	Ile	95	100	
Val	Asp	Lys	Tyr	Lys	Gly	Lys	Lys	Val	Asp	105	110	
Leu	Tyr	Gly	Ala	Tyr	Ala	Gly	Tyr	Gln	Cys	115	120	
Ala	Gly	Gly	Thr	Pro	Asn	Lys	Thr	Ala	Cys	125	130	
Met	Tyr	Gly	Gly	Val	Thr	Leu	His	Asp	Asn	135	140	
Asn	Arg	Leu	Thr	Glu	Glu	Lys	Lys	Val	Pro	145	150	
Ile	Asn	Leu	Trp	Leu	Asp	Gly	Lys	Gln	Asn	155	160	

Thr Val Pro Leu Glu Thr Val Lys Thr Asn	165	170
Lys Lys Asn Val Thr Val Gln Glu Leu Asp	175	180
Leu Gln Ala Arg Arg Tyr Leu Gln Glu Lys	185	190
Tyr Asn Leu Tyr Asn Ser Asp Val Phe Asp	195	200
Gly Lys Val Gln Arg Gly Leu Ile Val Phe	205	210
His Thr Ser Thr Glu Pro Ser Val Asn Tyr	215	220
Asp Leu Phe Gly Ala Gln Gly Gln Tyr Ser	225	230
Asn Thr Leu leu Arg Ile Tyr Arg Asp Asn	235	240
Lys Thr Ile Asn Ser Glu Asn Met His Ile	245	250
Asp Ile Tyr Leu Tyr Thr Ser	255	

(4) INFORMATION FOR SEQUENCE ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 757
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAGAAAAG CGAAGAAATA AATGAAAAAG ATTTGCGAAA	40
AAAGTCTGAA TTGCAGGGAA CAGCTTTAGG CAATCTTAAA	80
CAAATCTATT ATTACAATGA AAAAGCTAAA ACTGAAAATA	120
AAGAGAGTCA CGATCAATTT CGACAGCATA CTATATTGTT	160
TAAAGGCTTT TTTACAGATC ATTCGTGGTA TAACGATTTA	200

```

TTAGTACGTT TTGATTCAAA GGATATTGTT GATAAATATA 240
AAGGGAAAAA AGTAGACTTG TATGGTGCTT ATGCTGGTTA 280
TCAATGTGCG GGTGGTACAC CAAACAAAAC AGCTTGTATG 320
TATGGTGGTG TAACGTTACA TGATAATAAT CGATTGACCG 360
AAGAGAAAAA AGTGCCGATC AATTTATGGC TAGACGGTAA 400
ACAAAATACA GTACCTTTGG AACCGGTAA AACGAATAAG 440
AAAAATGTAA CTGTT CAGGA GTTGGATCTT CAAGCAAGAC 480
GTTATTTACA GGAAAAATAT AATTTATATA ACTCTGATGT 520
TTTTGATGGG AAGGTT CAGA GGGGATTAAT CGTGTTTCAT 560
ACTTCTACAG AACCTTCGGT TAATTACGAT TTATTTGGTG 600
CTCAAGGACA GTATTCAAAT AACTATTAA GAATATATAG 640
AGATAATAAA ACGATTA ACT CTGAAAACAT GCATATTGAT 680
ATATATTTAT ATACAAGTTA AACATGGTAG TTTTGACCAA 720
CGTAATGTTC AGATTATTAT GAACCGAGAA TAATCTA 757

```

(5) INFORMATION FOR SEQUENCE ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH:233
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Glu Lys Ser Glu Glu Ile Asn Glu Lys
      5                               10
Asp Leu Arg Lys Lys Ser Glu Leu Gln Gly
      15                             20
Thr Ala Leu Gly Asn Leu Lys Gln Ile Tyr
      25                             30
Tyr Tyr Asn Glu Lys Ala Lys Thr Glu Asn
      35                             40
Lys Glu Ser His Asp Gln Phe Arg Gln His
      45                             50

```

Thr	Ile	Leu	Phe	Lys	Gly	Phe	Phe	Thr	Asp	
				55					60	
His	Ser	Trp	Tyr	Asn	Asp	Leu	Leu	Val	Arg	
				65					70	
Phe	Asp	Ser	Lys	Asp	Ile	Val	Asp	Lys	Tyr	
				75					80	
Lys	Gly	Lys	Lys	Val	Asp	Leu	Tyr	Gly	Ala	
				85					90	
Tyr	Ala	Gly	Tyr	Gln	Cys	Ala	Gly	Gly	Thr	
				95					100	
Pro	Asn	Lys	Thr	Ala	Cys	Met	Tyr	Gly	Gly	
				105					110	
Val	Thr	Leu	His	Asp	Asn	Asn	Arg	Leu	Thr	
				115					120	
Glu	Glu	Lys	Lys	Val	Pro	Ile	Asn	Leu	Trp	
				125					130	
Leu	Asp	Gly	Lys	Gln	Asn	Thr	Val	Pro	Leu	
				135					140	
Glu	Thr	Val	Lys	Thr	Asn	Lys	Lys	Asn	Val	
				145					150	
Thr	Val	Gln	Glu	Leu	Asp	Leu	Gln	Ala	Arg	
				155					160	
Arg	Tyr	Leu	Gln	Glu	Lys	Tyr	Asn	Leu	Tyr	
				165					170	
Asn	Ser	Asp	Val	Phe	Asp	Gly	Lys	Val	Gln	
				175					180	
Arg	Gly	Leu	Ile	Val	Phe	His	Thr	Ser	Thr	
				185					190	
Glu	Pro	Ser	Val	Asn	Tyr	Asp	Leu	Phe	Gly	
				195					200	
Ala	Gln	Gly	Gln	Tyr	Ser	Asn	Thr	Leu	Leu	
				205					210	
Arg	Ile	Tyr	Arg	Asp	Asn	Lys	Thr	Ile	Asn	
				215					220	

Ser Glu Asn Met His Ile Asp Ile Tyr Leu
 225 230

Tyr Thr Ser

(6) INFORMATION FOR SEQUENCE ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1712
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAAC TAGGTA	GAAAAATAAT	TATGAGAAAA	CACTATGTTG	40
TTAAAGATGT	TTTCGTATAT	AAGTTTAGGT	GATGTATAGT	80
TACTTAATTT	TAAAAGCATA	ACTTAATTAA	TATAAATAAC	120
ATGAGATTAT	TAAATATAAT	TAAGTTTCTT	TTAATGTTTT	160
TTTAATTGAA	TATTTAAGAT	TATAACATAT	ATTTAAAGTG	200
TATCTAGATA	CTTTTTGGGA	ATGTTGGATA	AAGGAGATAA	240
AAAATGTATA	AGAGATTATT	TATTTACAT	GTAATTTTGA	280
TATTCGCACT	GATATTAGTT	ATTTCTACAC	CCAACGTTTT	320
AGCAGAGAGT	CAACCAGATC	CTAAACCAGA	TGAGTTGCAC	360
AAATCGAGTA	AATTCAC TGG	TTTGATGGAA	GATATGAAAG	400
TTTTGTATGA	TGATAATCAT	GTATCAGCAA	TAAACGTTAA	440
ATCTATAGAT	CAATTTCTAT	ACTTTGACTT	AATATATTCT	480
ATTAAGGACA	CTAAGTTAGG	GGATTATGAT	AATGTTTCGAG	520
TCGAATTTAA	AAACAAAGAT	TTAGCTGATA	AATACAAAGA	560
TAAATACGTA	GATGTGTTTG	GAGCTAATTA	TTATTATCAA	600
TGTTATTTTT	CTAAAAAAC	GAATGATATT	AATTCGCATC	640
AAACTGACAA	ACGAAAAACT	TGTATGTATG	GTGGTGTAAAC	680

TGAGCATAAT GGAAACCAAT TAGATAAATA TAGAAGTATT	720
ACTGTTTCGGG TATTTGAAGA TGGTAAAAAT TTATTATCTT	760
TTGACGTACA AACTAATAAG AAAAAGGTGA CTGCTCAAGA	800
ATTAGATTAC CTAACCTCGTC ACTATTTGGT GAAAAATAAA	840
AAACTCTATG AATTTAACAA CTCGCCTTAT GAAACGGGAT	880
ATATTAAATT TATAGAAAAT GAGAATAGCT TTTGGTATGA	920
CATGATGCCT GCACCAGGAG ATAAATTTGC CCAATCTAAA	960
TATTTAATGA TGTACAATGA CAATAAAATG GTTGATTCTA	1000
AAGATGTGAA GATTGAAGTT TATCTTACGA CAAAGAAAAA	1040
GTGAAATTAT ATTTTAGAAA AGTAAATATG AAGAGTTAGT	1080
AATTAAGGCA GGCACCTTATA GAGTACCTGC CTTTTCTAAT	1120
ATTATTTAGT TATAGTTATT TTTGTTATAT CTCTCTGATT	1160
TAGCATTAAC CCCTTGTTGC CATTATAGTT TTCACCAACT	1200
TTAGCTGAAA TTGGGGGATC ATTTTATCTT TTAATATGGA	1240
TAGTTACTGT GTCGCCGTTT TTAACGATTT GTTTCTCTTT	1280
TAATTTGTCA GTTAATTTTT TCCATGCATC ATTTGCGTCA	1320
AACCTATTTT CATTGATT TATTCTTGAC AAATCAATTC	1360
TTTTAACACT ATCGGTATTA ATCGGCTTGT TATTAAAATT	1400
ACTAAGTTCA TCTAAATCAG CTGTACCCGT AATACTACTT	1440
TCGCCACCAT TATTTAAATT GTACGTAACA CCAACTGTCT	1480
CATTGCTGT TTTATCGATA ATATTGCTT CTTTCAAAGC	1520
ATCTCTTACA TTTTCCATA AGTCTCTATC TGTATTTC	1560
GAAGCCTTTG CAACGTTATT AATACCATTA TAATTTGAAG	1600
AAGAATGAAA ACCTGAACCT ACTGTTGTTA AACTAAAGC	1640
ACTGCTATC AATGTTCTTG TTAATAGTTT TTTATTCATT	1680
TTATTTTCTC CTATAACTTA TTTGCAATCG AT	1712

(7) INFORMATION FOR SEQUENCE ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Tyr Lys Arg Leu Phe Ile Ser His Val
      5                               10

Ile Leu Ile Phe Ala Leu Ile Leu Val Ile
      15                             20

Ser Thr Pro Asn Val Leu Ala Glu Ser Gln
      25                             30

Pro Asp Pro Lys Pro Asp Glu Leu His Lys
      35                             40

Ser Ser Lys Phe Thr Gly Leu Met Glu Asp
      45                             50

Met Lys Val Leu Tyr Asp Asp Asn His Val
      55                             60

Ser Ala Ile Asn Val Lys Ser Ile Asp Gln
      65                             70

Phe Leu Tyr Phe Asp Leu Ile Tyr Ser Ile
      75                             80

Lys Asp Thr Lys Leu Gly Asp Tyr Asp Asn
      85                             90

Val Arg Val Glu Phe Lys Asn Lys Asp Leu
      95                             100

Ala Asp Lys Tyr Lys Asp Lys Tyr Val Asp
      105                            110

Val Phe Gly Ala Asn Tyr Tyr Tyr Gln Cys
      115                            120

Tyr Phe Ser Lys Lys Thr Asn Asp Ile Asn
      125                            130

Ser His Gln Thr Asp Lys Arg Lys Thr Cys
      135                            140

```

Met	Tyr	Gly	Gly	Val	Thr	Glu	His	Asn	Gly	
				145					150	
Asn	Gln	Leu	Asp	Lys	Tyr	Arg	Ser	Ile	Thr	
				155					160	
Val	Arg	Val	Phe	Glu	Asp	Gly	Lys	Asn	Leu	
				165					170	
Leu	Ser	Phe	Asp	Val	Gln	Thr	Asn	Lys	Lys	
				175					180	
Lys	Val	Thr	Ala	Gln	Glu	Leu	Asp	Tyr	Leu	
				185					190	
Thr	Arg	His	Tyr	Leu	Val	Lys	Asn	Lys	Lys	
				195					200	
Leu	Tyr	Glu	Phe	Asn	Asn	Ser	Pro	Tyr	Glu	
				205					210	
Thr	Gly	Tyr	Ile	Lys	Phe	Ile	Glu	Asn	Glu	
				215					220	
Asn	Ser	Phe	Trp	Tyr	Asp	Met	Met	Pro	Ala	
				225					230	
Pro	Gly	Asp	Lys	Phe	Ala	Gln	Ser	Lys	Tyr	
				235					240	
Leu	Met	Met	Tyr	Asn	Asp	Asn	Lys	Met	Val	
				245					250	
Asp	Ser	Lys	Asp	Val	Lys	Ile	Glu	Val	Tyr	
				255					260	
Leu	Thr	Thr	Lys	Lys	Lys					
				265						

(8) INFORMATION FOR SEQUENCE ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1712
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAACTAGGTA GAAAAATAAT TATGAGAAAA CACTATGTTG

TTAAAGATGT	TTTCGTATAT	AAGTTTAGGT	GATGTATAGT	80
TACTTAATTT	TAAAAGCATA	ACTTAATTAA	TATAAATAAC	120
ATGAGATTAT	TAAATATAAT	TAAGTTTCTT	TTAATGTTTT	160
TTTAATTGAA	TATTTAAGAT	TATAACATAT	ATTTAAAGTG	200
TATCTAGATA	CTTTTGGGA	ATGTTGGATA	AAGGAGATAA	240
AAAATGTATA	AGAGATTATT	TATTTACACAT	GTAATTTTGA	280
TATTCGCACT	GATATTAGTT	ATTTCTACAC	CCAACGTTTT	320
AGCAGAGAGT	CAACCAGATC	CTAAACCAGA	TGAGTTGCAC	360
AAATCGAGTA	AATTCACTGG	TTTGATGGAA	AATATGAAAG	400
TTTTGTATGA	TGATAATCAT	GTATCAGCAA	TAAACGTAA	440
ATCTATAGAT	CAATTCGAT	ACTTTGACTT	AATATATTCT	480
ATTAAGGACA	CTAAGTTAGG	GAATTATGAT	AATGTTTCGAG	520
TCGAATTTAA	AAACAAAGAT	TTAGCTGATA	AATACAAAGA	560
TAAATACGTA	GATGTGTTTG	GAGCTAATGC	TTATTATCAA	600
TGTGCTTTTT	CTAAAAAAC	GAATGATATT	AATTCGCATC	640
AAACTGACAA	ACGAAAAACT	TGTATGTATG	GTGGTGTAAC	680
TGAGCATAAT	GGAAACCAAT	TAGATAAATA	TAGAAGTATT	720
ACTGTTCTGGG	TATTTGAAGA	TGGTAAAAAT	TTATTATCTT	760
TTGACGTACA	AACTAATAAG	AAAAAGGTGA	CTGCTCAAGA	800
ATTAGATTAC	CTAACTCGTC	ACTATTTGGT	GAAAAATAAA	840
AAACTCTATG	AATTTAACAA	CTCGCCTTAT	GAAACGGGAT	880
ATATTAAATT	TATAGAAAAT	GAGAATAGCT	TTTGGTATGA	920
CATGATGCCT	GCACCAGGAG	ATAAATTGA	CCAATCTAAA	960
TATTTAATGA	TGTACAATGA	CAATAAAATG	GTTGATTCTA	1000
AAGATGTGAA	GATTGAAGTT	TATCTTACGA	CAAAGAAAAA	1040
GTGAAATTAT	ATTTTAGAAA	AGTAAATATG	AAGAGTTAGT	1080
AATTAAGGCA	GGCACTTATA	GAGTACCTGC	CTTTTCTAAT	1120

```

ATTATTTAGT TATAGTTATT TTTGTTATAT CTCTCTGATT 1160
TAGCATTAAC CCCTTGTTGC CATTATAGTT TTCACCAACT 1200
TTAGCTGAAA TTGGGGGATC ATTTTATCT TTAATATGGA 1240
TAGTTACTGT GTCGCCGTTT TTAACGATTT GTTCTCTTT 1280
TAATTTGTCA GTTAATTTTT TCCATGCATC ATTTGCGTCA 1320
AACCTATTTT CATTGATGATT TATTCTTGAC AAATCAATTC 1360
TTTAAACACT ATCGGTATTA ATCGGCTTGT TATTAAAATT 1400
ACTAAGTTCA TCTAAATCAG CTGTACCCGT AATACTACTT 1440
TCGCCACCAT TATTTAAATT GTACGTAACA CCAACTGTCT 1480
CATTGCTGT TTTATCGATA ATATTGCTT CTTCAAAGC 1520
ATCTCTTACA TTTTCCATA AGTCTCTATC TGTATTTC 1560
GAAGCCTTTG CAACGTTATT AATACCATTA TAATTTGAAG 1600
AAGAATGAAA ACCTGAACCT ACTGTTGTTA AACTAAAGC 1640
ACTTGCTATC AATGTTCTTG TTAATAGTTT TTTATTCATT 1680
TTATTTTCTC CTATAACTTA TTTGCAATCG AT 1712

```

(9) INFORMATION FOR SEQUENCE ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Tyr Lys Arg Leu Phe Ile Ser His Val
      5                               10

Ile Leu Ile Phe Ala Leu Ile Leu Val Ile
      15                             20

Ser Thr Pro Asn Val Leu Ala Glu Ser Gln
      25                             30

```

Pro	Asp	Pro	Lys	Pro	Asp	Glu	Leu	His	Lys	
				35					40	
Ser	Ser	Lys	Phe	Thr	Gly	Leu	Met	Glu	Asn	
				45					50	
Met	Lys	Val	Leu	Tyr	Asp	Asp	Asn	His	Val	
				55					60	
Ser	Ala	Ile	Asn	Val	Lys	Ser	Ile	Asp	Gln	
				65					70	
Phe	Arg	Tyr	Phe	Asp	Leu	Ile	Tyr	Ser	Ile	
				75					80	
Lys	Asp	Thr	Lys	Leu	Gly	Asn	Tyr	Asp	Asn	
				85					90	
Val	Arg	Val	Glu	Phe	Lys	Asn	Lys	Asp	Leu	
				95					100	
Ala	Asp	Lys	Tyr	Lys	Asp	Lys	Tyr	Val	Asp	
				105					110	
Val	Phe	Gly	Ala	Asn	Ala	Tyr	Tyr	Gln	Cys	
				115					120	
Ala	Phe	Ser	Lys	Lys	Thr	Asn	Asp	Ile	Asn	
				125					130	
Ser	His	Gln	Thr	Asp	Lys	Arg	Lys	Thr	Cys	
				135					140	
Met	Tyr	Gly	Gly	Val	Thr	Glu	His	Asn	Gly	
				145					150	
Asn	Gln	Leu	Asp	Lys	Tyr	Arg	Ser	Ile	Thr	
				155					160	
Val	Arg	Val	Phe	Glu	Asp	Gly	Lys	Asn	Leu	
				165					170	
Leu	Ser	Phe	Asp	Val	Gln	Tyr	Asn	Lys	Lys	
				175					180	
Lys	Val	Thr	Ala	Gln	Glu	Leu	Asp	Tyr	Leu	
				185					190	
Thr	Arg	His	Tyr	Leu	Val	Lys	Asn	Lys	Lys	
				195					200	
Leu	Tyr	Glu	Phe	Asn	Asn	Ser	Pro	Tyr	Glu	
				205					210	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300

Thr	Gly	Tyr	Ile	Lys	Phe	Ile	Glu	Asn	Glu	
				215					220	
Asn	Ser	Phe	Trp	Tyr	Asp	Met	Met	Pro	Ala	
				225					230	
Pro	Gly	Asp	Lys	Phe	Asp	Gln	Ser	Lys	Tyr	
				235					240	
Leu	Met	Met	Tyr	Asn	Asp	Asn	Lys	Met	Val	
				245					250	
Asp	Ser	Lys	Asp	Val	Lys	Ile	Glu	Val	Tyr	
				255					260	
Leu	Thr	Thr	Lys	Lys	Lys					
				265						

(10) INFORMATION FOR SEQUENCE ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1388
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGAGTCAAC CAGATCCTAA ACCAGATGAG TTGCACAAAT	40
CGAGTAAATT CACTGGTTTG ATGGAAAATA TGAAAGTTTT	80
GTATGATGAT AATCATGTAT CAGCAATAAA CGTTAAATCT	120
ATAGATCAAT TTCGATACTT TGACTTAATA TATTCTATTA	160
AGGACACTAA GTTAGGGAAT TATGATAATG TTCGAGTCGA	200
ATTTAAAAAC AAAGATTTAG CTGATAAATA CAAAGATAAA	240
TACGTAGATG TGTTTGAGGC TAATGCTTAT TATCAATGTG	280
CTTTTCTTAA AAAACGAAT GATATTAATT CGCATCAAAC	320
TGACAAACGA AAAACTTGTA TGTATGGTGG TGTAAC TGAG	360
CATAATGGAA ACCAATTAGA TAAATATAGA AGTATTACTG	400
TTCGGGTATT TGAAGATGGT AAAAATTTAT TATCTTTTGA	440

CGTACAACT AATAAGAAAA AGGTGACTGC TCAAGAATTA	480
GATTACCTAA CTCGTCCTA TTTGGTGAAA AATAAAAAAC	520
TCTATGAATT TAACAACTCG CCTTATGAAA CGGGATATAT	560
TAAATTTATA GAAAATGAGA ATAGCTTTTG GTATGACATG	600
ATGCCTGCAC CAGGAGATAA ATTTGACCAA TCTAAATATT	640
TAATGATGTA CAATGACAAT AAAATGGTTG ATTCTAAAGA	680
TGTGAAGATT GAAGTTTATC TTACGACAAA GAAAAAGTGA	720
AATTATATTT TAGAAAAGTA AATATGAAGA GTTAGTAATT	760
AAGGCAGGCA CTTATAGAGT ACCTGCCTTT TCTAATATTA	800
TTTAGTTATA GTTATTTTTG TTATATCTCT CTGATTTAGC	840
ATTAACCCCT TGTTGCCATT ATAGTTTTCAC CAACTTTAG	880
CTGAAATTGG GGGATCATTT TTATCTTTAC TATGGATAGT	920
TACTGTGTCG CCGTTTTTAA CGATTTGTTT CTCTTTTAAT	960
TTGTCAGTTA ATTTTTTCCA TGCATCATTT GCGTCAAACC	1000
TATTTCCATT TGGATTTATT CTTGACAAAT CAATTCTTTT	1040
AACACTATCG GTATTAATCG GCTTGTTATT AAAATTACTA	1080
AGTTCATCTA AATCAGCTGT ACCCGTAATA CTACTTTCGC	1120
CACCATTATT TAAATTGTAC GTAACACCAA CTGTCTCATT	1160
TGCTGTTTTA TCGATAATAT TTGCTTCTTT CAAAGCATCT	1200
CTTACATTTT TCCATAAGTC TCTATCTGTT ATTTCAGAAG	1240
CCTTTGCAAC GTTATTAATA CCATTATAAT TTGAAGAAGA	1280
ATGAAAACCT GAACCTACTG TTGTTAAAAC TAAAGCACTT	1320
GCTATCAATG TTCTTGTTAA TAGTTTTTTA TTCATTTTAT	1360
TTTCTCCTAT AACTTATTTG CAATCGAT	1388

(11) INFORMATION FOR SEQUENCE ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 239

(B) TYPE: Amino Acid
 (C) STRANDEDNESS: Unknown
 (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ser	Gln	Pro	Asp	Pro	Lys	Pro	Asp	Glu	5	10
Leu	His	Lys	Ser	Ser	Lys	Phe	Thr	Gly	Leu	15	20
Met	Glu	Asn	Met	Lys	Val	Leu	Tyr	Asp	Asp	25	30
Asn	His	Val	Ser	Ala	Ile	Asn	Val	Lys	Ser	35	40
Ile	Asp	Gln	Phe	Arg	Tyr	Phe	Asp	Leu	Ile	45	50
Tyr	Ser	Ile	Lys	Asp	Thr	Lys	Leu	Gly	Asn	55	60
Tyr	Asp	Asn	Val	Arg	Val	Glu	Phe	Lys	Asn	65	70
Lys	Asp	Leu	Ala	Asp	Lys	Tyr	Lys	Asp	Lys	75	80
Tyr	Val	Asp	Val	Phe	Gly	Ala	Asn	Ala	Tyr	85	90
Tyr	Gln	Cys	Ala	Phe	Ser	Lys	Lys	Thr	Asn	95	100
Asp	Ile	Asn	Ser	His	Gln	Thr	Asp	Lys	Arg	105	110
Lys	Thr	Cys	Met	Tyr	Gly	Gly	Val	Thr	Glu	115	120
His	Asn	Gly	Asn	Gln	Leu	Asp	Lys	Tyr	Arg	125	130
Ser	Ile	Thr	Val	Arg	Val	Phe	Glu	Asp	Gly	135	140
Lys	Asn	Leu	Leu	Ser	Phe	Asp	Val	Gln	Thr	145	150
Asn	Lys	Lys	Lys	Val	Thr	Ala	Gln	Glu	Leu		

	155		160
Asp Tyr Leu Thr	Arg His Tyr Leu Val	Lys	
	165		170
Asn Lys Lys Leu	Tyr Glu Phe Asn Asn	Ser	
	175		180
Pro Tyr Glu Thr	Gly Tyr Ile Lys Phe	Ile	
	185		190
Glu Asn Glu Asn	Ser Phe Trp Tyr Asp	Met	
	195		200
Met Pro Ala Pro	Gly Asp Lys Phe Asp	Gln	
	205		210
Ser Lys Tyr Leu	Met Met Tyr Asn Asp	Asn	
	215		220
Lys Met Val Asp	Ser Lys Asp Val Lys	Ile	
	225		230
Glu Val Tyr Leu	Thr Thr Lys Lys Lys		
	235		

(12) INFORMATION FOR SEQUENCE ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 731
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TAAGGAGAAT TAAAAATGAA TAAAAAATTA CTAATGAATT	40
TTTTTATCGT AAGCCCTTTG TTGCTTGCGA CAACTGCTAC	80
AGATTTTACC CCTGTTCCCT TATCATCTAA TCAAATAATC	120
AAAACTGCAA AAGCATCTAC AAACGATAAT ATAAAGGATT	160
TGCTAGACTG GTATAGTAGT GGGTCTGACA CTTTACAAA	200
TAGTGAAGTT TTAGATAATT CCAGAGGATC TATGCGTATA	240
AAAAACACAG ATGGCAGCAT CAGCTTGATA ATTTTCCGA	280

```

GTCCTTATTA TAGCCCTGCT TTTACAAAAG GGGAAAAAGT 320
TGACTTAAAC ACAAAAAGAA CTAAAAAAG CCAACATACT 360
AGCGAAGGAA CTTATATCCA TTTCCAAATA AGTGGCGTTA 400
CAAATACTGA AAAATTACCT ACTCCAATAG AACTACCTTT 440
AAAAGTTAAG GTTCATGGTA AAGATAGCCC CTAAAGTAT 480
GGGCCAAAGT TCGATAAAAA ACAATTAGCT ATATCAACTT 520
TAGACTTTGA AATTTCGTCAT CAGCTAACTC AAATACATGG 560
ATTATATCGT TCAAGCGATA AAACGGGTGG TTATTGGAAA 600
ATAACAATGA ATGACGGATC CACATATCAA AGTGATTTAT 640
CTAAAAAGTT TGAATACAAT ACTGAAAAAC CACCTATAAA 680
TATTGATGAA ATAAAACTA TAGAAGCAGA AATTAATTAA 720
TTTACCACTT T 731

```

(13) INFORMATION FOR SEQUENCE ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Asn Lys Lys Leu Leu Met Asn Phe Phe
                    5                      10
Ile Val Ser Pro Leu Leu Leu Ala Thr Thr
                    15                      20
Ala Thr Asp Phe Thr Pro Val Pro Leu Ser
                    25                      30
Ser Asn Gln Ile Ile Lys Thr Ala Lys Ala
                    35                      40
Ser Thr Asn Asp Asn Ile Lys Asp Leu Leu
                    45                      50
Asp Trp Tyr Ser Ser Gly Ser Asp Thr Phe
                    55                      60

```

Thr	Asn	Ser	Glu	Val	Leu	Asp	Asn	Ser	Arg		
				65					70		
Gly	Ser	Met	Arg	Ile	Lys	Asn	Thr	Asp	Gly		
				75					80		
Ser	Ile	Ser	Leu	Ile	Ile	Phe	Pro	Ser	Pro		
				85					90		
Tyr	Tyr	Ser	Pro	Ala	Phe	Thr	Lys	Gly	Glu		
				95					100		
Lys	Val	Asp	Leu	Asn	Thr	Lys	Arg	Thr	Lys		
				105					110		
Lys	Ser	Gln	His	Thr	Ser	Glu	Gly	Thr	Tyr		
				115					120		
Ile	His	Phe	Gln	Ile	Ser	Gly	Val	Thr	Asn		
				125					130		
Thr	Glu	Lys	Leu	Pro	Thr	Pro	Ile	Glu	Leu		
				135					140		
Pro	Leu	Lys	Val	Lys	Val	His	Gly	Lys	Asp		
				145					150		
Ser	Pro	Leu	Lys	Tyr	Gly	Pro	Lys	Phe	Asp		
				155					160		
Lys	Lys	Gln	Leu	Ala	Ile	Ser	Thr	Leu	Asp		
				165					170		
Phe	Glu	Ile	Arg	His	Gln	Leu	Thr	Gln	Ile		
				175					180		
His	Gly	Leu	Tyr	Arg	Ser	Ser	Asp	Lys	Thr		
				185					190		
Gly	Gly	Tyr	Trp	Lys	Ile	Thr	Met	Asn	Asp		
				195					200		
Gly	Ser	Thr	Tyr	Gln	Ser	Asp	Leu	Ser	Lys		
				205					210		
Lys	Phe	Glu	Tyr	Asn	Thr	Glu	Lys	Pro	Pro		
				215					220		
Ile	Asn	Ile	Asp	Glu	Ile	Lys	Thr	Ile	Glu		
				225					230		
Ala	Glu	Ile	Asn								

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

(14) INFORMATION FOR SEQUENCE ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1095
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```

ATCATTAAT ATAATTAATT TTCTTTTAAT ATTTTTTTTAA      40
TTGAATATTT AAGATTATAA GATATATTTA AAGTGTATCT      80
AGATACTTTT TGGGAATGTT GGATGAAGGA GATAAAAATG     120
AATAAGAGTC GATTTATTTT ATGCGTAATT TTGATATTCG     160
CACTTATACT AGTTCTTTTT ACACCCAACG TATTAGCAGA     200
GAGCCAACCA GACCCTACGC CAGATGAGTT GCACAAAGCG     240
AGTAAATTCA CTGGTTTGAT GGAAAATATG AAAGTTTAT      280
ATGATGATCA TTATGTATCA GCAACTAAAG TTAAGTCTGT      320
AGATAAATTT AGGGCACATG ATTTAATTTA TAACATTAGT      360
GATAAAAAAC TGAAAATTA TGACAAAGTG AAAACAGAGT      400
TATTAAATGA AGGTTTAGCA AAGAAGTACA AAGATGAAGT      440
AGTTGATGTG TATGGATCAA ATTACTATGT AAAGTGCTAT      480
TTTTCATCCA AAGATAATGT AGGTAAAGTT ACAGGTGGCA      520
AAACTTGATG GTATGGAGGA ATAACAAAAC ATGAAGGAAA      560
CCACTTTGAT AATGGGAACT TACAAAATGT ACTTATAAGA      600
GTTTATGAAA ATAAAAGAAA CACAATTTCT TTTGAAGTGC      640
AAACTGATAA GAAAAGTGTA ACAGCTCAAG AACTAGACAT      680
AAAAGCTAGG AATTTTTTTAA TTAATAAAAA AAATTTGTAT      720
GAGTTTAACA GTTCACCATA TGAAACAGGA TATATAAAAT      760
TTATTGAAAA TAACGGCAAT ACTTTTGGT ATGATATGAT      800

```

```

GCCTGCACCA GCGATAAGT TTGACCAATC TAAATATTTA      840
ATGATGTACA ACGACAATAA AACGGTTGAT TCTAAAAGTG      880
TGAAGATAGA AGTCCACCTT ACAACAAAGA ATGGATAATG      920
TTAATCCGAT TTTGATATAA AAAGTGAAAG TATTAGATAT      960
ATTTGAAAGG TAAGTACTTC GGTGCTTGCC TTTTtaggat     1000
GCATATATAT AGATTAAACC GCACTTCTAT ATTAATAGAA      1040
AGTGCGGTTA TTTATACACT CAATCTAAAC TATAATAATT      1080
GGAATCATCT TCAAA                                  1095

```

(15) INFORMATION FOR SEQUENCE ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Asn Lys Ser Arg Phe Ile Ser Cys Val
                    5                      10
Ile Leu Ile Phe Ala Leu Ile Leu Val Leu
                    15                      20
Phe Thr Pro Asn Val Leu Ala Glu Ser Gln
                    25                      30
Pro Asp Pro Thr Pro Asp Glu Leu His Lys
                    35                      40
Ala Ser Lys Phe Thr Gly Leu Met Glu Asn
                    45                      50
Met Lys Val Leu Tyr Asp Asp His Tyr Val
                    55                      60
Ser Ala Thr Lys Val Lys Ser Val Asp Lys
                    65                      70
Phe Arg Ala His Asp Leu Ile Tyr Asn Ile
                    75                      80

```

Ser	Asp	Lys	Lys	Leu	Lys	Asn	Tyr	Asp	Lys	
				85					90	
Val	Lys	Thr	Glu	Leu	Leu	Asn	Glu	Gly	Leu	
				95					100	
Ala	Lys	Lys	Tyr	Lys	Asp	Glu	Val	Val	Asp	
				105					110	
Val	Tyr	Gly	Ser	Asn	Tyr	Tyr	Val	Asn	Cys	
				115					120	
Tyr	Phe	Ser	Ser	Lys	Asp	Asn	Val	Gly	Lys	
				125					130	
Val	Thr	Gly	Gly	Lys	Thr	Cys	Met	Tyr	Gly	
				135					140	
Gly	Ile	Thr	Lys	His	Glu	Gly	Asn	His	Phe	
				145					150	
Asp	Asn	Gly	Asn	Leu	Gln	Asn	Val	Leu	Ile	
				155					160	
Arg	Val	Tyr	Glu	Asn	Lys	Arg	Asn	Thr	Ile	
				165					170	
Ser	Phe	Glu	Val	Gln	Thr	Asp	Lys	Lys	Ser	
				175					180	
Val	Thr	Ala	Gln	Glu	Leu	Asp	Ile	Lys	Ala	
				185					190	
Arg	Asn	Phe	Leu	Ile	Asn	Lys	Lys	Asn	Leu	
				195					200	
Tyr	Glu	Phe	Asn	Ser	Ser	Phe	Tyr	Glu	Thr	
				205					210	
Gly	Tyr	Ile	Lys	Phe	Ile	Glu	Asn	Asn	Gly	
				215					220	
Asn	Thr	Phe	Trp	Tyr	Asp	Met	Met	Pro	Ala	
				225					230	
Pro	Gly	Asp	Lys	Phe	Asp	Gln	Ser	Lys	Tyr	
				235					240	
Leu	Met	Met	Tyr	Asn	Asp	Asn	Lys	Thr	Val	
				245					250	
Asp	Ser	Lys	Ser	Val	Lys	Ile	Glu	Val	His	
				255					260	

Leu Thr Thr Lys Asn Gly
265

(16) INFORMATION FOR SEQUENCE ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1837
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TCATGTTTGA CAGCTTATCA TCGATAAGCT TACTTTTCGA	40
ATCAGGTCTA TCCTTGAAAC AGGTGCAACA TAGATTAGGG	80
CATGGAGATT TACCAGACAA CTATGAACGT ATATACTCAC	120
ATCACGCAAT CGGCAATTGA TGACATTGGA ACTAAATTCA	160
ATCAATTTGT TACTAACAAG CAACTAGATT GACAACTAAT	200
TCTCAACAAA CGTTAATTTA ACAACATTCA AGTAACTCCC	240
ACCAGCTCCA TCAATGCTTA CCGTAAGTAA TCATAACTTA	280
CTAAACCTT GTTACATCAA GGTTTTTTCT TTTTGTCTTG	320
TTCATGAGTT ACCATAACTT TCTATATTAT TGACAACTAA	360
ATTGACAACT CTTCAATTAT TTTTCTGTCT ACTCAAAGTT	400
TTCTTCATTT GATATAGTCT AATTCACCA TCACTTCTTC	440
CACTCTCTCT ACCGTCACAA CTTCATCATC TCTCACTTTT	480
TCGTGTGGTA ACACATAATC AAATATCTTT CCGTTTTTAC	520
GCACTATCGC TACTGTGTCA CCTAAAATAT ACCCCTTATC	560
AATCGCTTCT TTAAACTCAT CTATATATAA CATATTTTCAT	600
CCTCCTACCT ATCTATTCGT AAAAAGATAA AAATAACTAT	640
TGTTTTTTTT GTTATTTTAT AATAAAATTA TTAATATAAG	680
TTAATGTTTT TTAAAAATAT ACAATTTTAT TCTATTTATA	720

GTTAGCTATT	TTTTCATTGT	TAGTAATATT	GGTGAATTGT	760
AATAACCTTT	TTAAATCTAG	AGGAGAACCC	AGATATAAAA	800
TGGAGGAATA	TTAATGGAAA	ACAATAAAAA	AGTATTGAAG	840
AAAATGGTAT	TTTTTGTTTT	AGTGACATTT	CTTGGACTION	880
CAATCTCGCA	AGAGGTATTT	GCTCAACAAG	ACCCCGATCC	920
AAGCCAACTT	CACAGATCTA	GTTTAGTTAA	AAACCTTCAA	960
AATATATATT	TTCTTTATGA	GGGTGACCCT	GTTACTCACG	1000
AGAATGTGAA	ATCTGTTGAT	CAACTTAGAT	CTCACGATTT	1040
AATATATAAT	GTTTCAGGGC	CAAATTATGA	TAAATTAAAA	1080
ACTGAACTTA	AGAACCAAGA	GATGGCAACT	TTATTTAAGG	1120
ATAAAAACGT	TGATATTTAT	GGTGTAGAAT	ATTACCATCT	1160
CTGTTATTTA	TGTGAAAATG	CAGAAAGGAG	TGCATGTATC	1200
TACGGAGGGG	TAACAAATCA	TGAAGGGAAT	CATTTAGAAA	1240
TTCCATAAAA	GATAGTCGTT	AAAGTATCAA	TCGATGGTAT	1280
CCAAAGCCTA	TCATTTGATA	TTGAAACAAA	TAAAAAAATG	1320
GTAAGTCTC	AAGAATTAGA	CTATAAAGTT	AGAAAAATATC	1360
TTACAGATAA	TAAGCAACTA	TATACTAATG	GACCTTCTAA	1400
ATATGAAACT	GGATATATAA	AGTTCATACC	TAAGAATAAA	1440
GAAAGTTTTT	GGTTTGATTT	TTTCCCTGAA	CCAGAATTTA	1480
CTCAATCTAA	ATATCTTATG	ATATATAAAG	ATAATGAAAC	1520
GCTTGACTCA	AACACAAGCC	AAATTGAAGT	CTACCTAACA	1560
ACCAAGTAAC	TTTTTGCTTT	TGGCAACCTT	ACCTACTGCT	1600
GGATTTAGAA	ATTTTATTGC	AATTCTTTTA	TTAATGTAAA	1640
AACCGCTCAT	TTGATGAGCG	GTTTTGTCTT	ATCTAAAGGA	1680
GCTTTACCTC	CTAATGCTGC	AAAATTTTAA	ATGTTGGATT	1720
TTTGTATTTG	TCTATTGTAT	TTGATGGGTA	ATCCCATTTT	1760
TCGACAGACA	TCGTCGTGCC	ACCTCTAACA	CCAAAATCAT	1800

AGACAGGAGC TTGTAGCTTA GCAACTATTT TATCGTC

1837

(17) INFORMATION FOR SEQUENCE ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Unknown
- (D) TOPOLOGY: Unknown

(ii) Molecule type: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Glu Asn Asn Lys Lys Val Leu Lys Lys
5 10

Met Val Phe Phe Val Leu Val Thr Phe Leu
15 20

Gly Leu Thr Ile Ser Gln Glu Val Phe Ala
25 30

Gln Gln Asp Pro Asp Pro Ser Gln Leu His
35 40

Arg Ser Ser Leu Val Lys Asn Leu Gln Asn
45 50

Ile Tyr Phe Leu Tyr Glu Gly Asp Pro Val
55 60

Thr His Glu Asn Val Lys Ser Val Asp Gln
65 70

Leu Arg Ser His Asp Leu Ile Tyr Asn Val
75 80

Ser Gly Pro Asn Tyr Asp Lys Leu Lys Thr
85 90

Glu Leu Lys Asn Gln Glu Met Ala Thr Leu
95 100

Phe Lys Asp Lys Asn Val Asp Ile Tyr Gly
105 110

Val Glu Tyr Tyr His Leu Cys Tyr Leu Cys
115 120

Glu Asn Ala Glu Arg Ser Ala Cys Ile Tyr
125 130

Gly Gly Val Thr Asn His Glu Gly Asn His

	135	140
Leu Glu Ile Pro Lys Lys Ile Val Val Lys	145	150
Val Ser Ile Asp Gly Ile Gln Ser Leu Ser	155	160
Phe Asp Ile Glu Thr Asn Lys Lys Met Val	165	170
Thr Ala Gln Glu Leu Asp Tyr Lys Val Arg	175	180
Lys Tyr Leu Thr Asp Asn Lys Gln Leu Tyr	185	190
Thr Asn Gly Pro Ser Lys Tyr Glu Thr Gly	195	200
Tyr Ile Lys Phe Ile Pro Lys Asn Lys Glu	205	210
Ser Phe Trp Phe Asp Phe Phe Pro Glu Pro	215	220
Glu Phe Thr Gln Ser Lys Tyr Leu Met Ile	225	230
Tyr Lys Asp Asn Glu Thr Leu Asp Ser Asn	235	240
Thr Ser Gln Ile Glu Val Tyr Leu Thr Thr	245	250
Lys		

SEQ ID NO:17**sense primer for cloning SpeA**

5' CTCG CAA GAG GTA CAT ATG CAA CAA GAC 3' (SEQ ID NO:),
 sense primer to introduce a unique NdeI site;

SEQ ID NO:18**antisense primer for cloning SpeA**

5' GCA GTA GGT AAG CTT GCC AAA AGC 3' (SEQ ID NO:)

SEQ ID NO:19**1. SpeA forward primer, including NdeI site:**

5' GATATACATATGCAACAAGACCCCGATCCAAGCC 3' 34-mer
 :

SEQ ID NO:20

SpeA reverse primer, adds SpeB overlap

5' GAGATTTAACAACCTGGTTGCTTGGTTGTTAGGTAGAC 3' 37-mer

SEQ ID NO:21

3.SpeB forward primer, adds SpeA overlap:

5' GTCTACCTAACAACCAAGC A A C C A G T T G T T A A A T C T C 3' 37-mer

SEQ ID NO:22

4. SpeB reverse primer; adds stop site (Amber) and maintains BamHI site:

5' GAATTCGGATCCGCTAGCCTACAACAG 3' 27-mer

SEQ ID NO:23

SpeA(L42R)-SpeB(C47S) gene insert DNA sequence

1
atgcaacaag accccgatcc aagccaactt cacagatcta gtttagtta
51
aaaccttcaa aatatatatt ttctttatga gggtgaccct gttactcacg
101
agaatgtgaa atctgttgat caacttcgat ctcacgattt aatatataat
151
gtttcagggc caaattatga taaattaaaa actgaactta agaaccaaga
201
gatggcaact ttatttaagg ataaaaacat tgatatttat ggtgtagaat
251
attaccatct ctgttattta tgtgaaaatg cagaaaggag tgcattgtatc
301
tacggagggg taacaaatcg tgaagggaat catttagaaa ttcctaaaaa
351
gatagtcgtc aaagtatcaa tcgatgggat acaaagccta tcatttgata
401
ttgaaacaaa taaaaaatg gtaactgctc aagaattaga ctataaagtt
451
agaaaatatc ttacagataa taagcaacta tataactaatg gaccttctaa
501
atatgaaact ggatatataa agttcatacc taagaataaa gaaagttttt
551
ggtttgattt tttccctgaa ccagaattta ctcaatctaa atatcttatg
601
atatataaag ataatgaaac gcttgactca aacacaagcc aaattgaagt
651
ctacctaaca accaagcaac cagttggtta atctctcctt gattcaaaag
701
gcattcatta caatcaaggt aacccttaca acctattgac acctgttatt
751
gaaaaagtaa aaccaggtga acaatctttt gtaggtcaac atgcagctac
801
aggatgtgtt gctactgcaa ctgctcaa at tatgaaatat cataattacc
851

ctaacaaagg gttgaaagac tacacttaca cactaagctc aaataaccca
 901
 tatttcaacc atcctaagaa cttgtttgca gctatctcta ctagacaata
 951
 caactggaac aacatcctac ctacttatag cggaagagaa tctaacgttc
 1001
 aaaaaatggc gatttcagaa ttgatggctg atgttggtat ttcagtagac
 1051
 atggattatg gtccatctag tggttctgca ggtagctctc gtgttcaaag
 1101
 agccttgaaa gaaaactttg gctacaacca atctgttcac caaatcaacc
 1151
 gtagcgactt tagcaaacia gattgggaag cacaaattga caaagaatta
 1201
 tctcaaaacc aaccagtata ctaccaaggt gtcggtaaag taggcggaca
 1251
 tgcctttgtt atcgatggtg ctgacggacg taacttctac catgttaact
 1301
 ggggttgggg tggagtctct gacggcttct tccgtcttga cgcactaaac
 1351
 ccttcagctc ttggtactgg tggcggcgca ggcggcttca acggttacca
 1401
 aagtgcgtgt gtaggctag

SEQ ID NO:24

Full-length SpeB polypeptide (Kagawa et al., PNAS 97:2235-2240. 2000):

1	MNKKKL	GIRL	LSLLAL	GGFV	LANPVF	ADQN	FARNEKE	EAKD	SAITFIQ	KSA	AIKAGAR	SAE
61	DIKLDK	VNLG	GELSGS	NMYV	YNISTG	GFVI	VSGDKR	SPEI	LGYSTS	GSFD	ANGKENI	ASF
121	MESYVE	QIKE	NKKLDT	TYAG	TAEIKQ	PVVK	SLDSKG	IIHY	NQGNPY	NLLT	PVIEKVK	PGE
181	QSFVQG	HAAT	GCVATA	TATQI	MKYHNP	PNKG	LKDYTY	TLSS	NNPYFN	HHPKN	LFAAIST	RQY
241	NWNNIL	PTYS	GRESNV	QKMA	ISELMAD	VGI	SVDMDY	GPSS	GSAGSS	RVQR	ALKENFG	YNQ
301	SVHQIN	RGDF	SKQDWE	AQID	KELSQN	QPVY	YQGVGK	VGGH	AFVIDG	ADGR	NFYHVN	WGWG
361	GVSDGF	FRLD	ALNPSA	LGTG	GGAGGF	NGYQ	SAVVGI	KP				

SEQ ID NO:25

SpeB clone used for fusion, mature polypeptide. Estimated M_r = 28.75 kDa

```

1 QPVVKSLLDS KGIHYNQGNP YNLLTPVIEK VKPGEQSFVG QHAATGCVAT
51 ATAQIMKYHN YPNKGLKDYT YTLSSNNPYF NHPKNLF AAI STRQYNWNNI
101 LPTYSGRESN VQKMAISELM ADVGISVDMD YGPSSGSAGS SRVQRALKEN
151 FGYNQSVHQI NRSDFSQDWE AQIDKELSON QPVYYQGGKV GGHAFFVIDGA
201 DGRNFYHVNW GWGGVSDGFF RLDALNPSAL GTGGGAGGFN GYQSAVVG

```

SEQ ID NO: 26

L42R SpeA mutant clone used for fusion. Estimated M_r = 25.2 kDa

```

1 MQQDPDP SQL HRSSLVKNLQ NIYFLYEGDP VTHENVK SVD QLRSHDLIYN
51 VSGPNYDKLK TELKNQEMAT LFKDKNIDIY GVEYYHLCYL CENAERSACI
101 GGVNREGNH LEIPKKIVVK VSIDGIQSL S FDIETNKKMV TAQELDYKVR
151 KYLTDNKQLY TNGPSKYETG YIKFIPKNKE SFWFDFFP EP EFTQSKYLM I
201 YKDNETLDSN TQIEVYLTTK

```

SEQ ID NO:27

SpeA [L42R]-SpeB [C47S] fusion. Estimated M_r = 54 kDa

```

1 MQQDPDP SQL HRSSLVKNLQ NIYFLYEGDP VTHENVK SVD QLRSHDLIYN
51 VSGPNYDKLK TELKNQEMAT LFKDKNIDIY GVEYYHLCYL CENAERSACI
101 GGVNREGNH LEIPKKIVVK VSIDGIQSL S FDIETNKKMV TAQELDYKVR
151 KYLTDNKQLY TNGPSKYETG YIKFIPKNKE SFWFDFFP EP EFTQSKYLM I
201 YKDNETLDSN TQIEVYLTTK QPVVKSLLDS KGIHYNQGNP YNLLTPVIEK
251 VKPGEQSFVG QHAATGCVAT ATAQIMKYHN YPNKGLKDYT YTLSSNNPYF
301 NHPKNLF AAI STRQYNWNNI LPTYSGRESN VQKMAISELM ADVGISVDMD
351 YGPSSGSAGS SRVQRALKEN FGYNQSVHQI NRSDFSQDWE AQIDKELSON
401 QPVYYQGGKV GGHAFFVIDGA DGRNFYHVNW GWGGVSDGFF RLDALNPSAL
551 GTGGGAGGFN GYQSAVVG

```

SEQ ID NO:28

1. SpeA forward primer, including NdeI site:

5' GATATACATATGCAACAAGACCCCGATCCAAGCC 3' 34-mer

SEQ ID NO:29

2. SpeA reverse primer; kills NdeI site, adds SpeB overlap:

5' CATGTGTATATCTCCTTCCTTGGTTGTTAGGTAGAC 3' 36-mer

SEQ ID NO:30

3. SpeB forward primer; kills NdeI site, adds SpeA overlap:

5' GTCTACCTAACAACCAAGGAAGGAGATATACACATG 3' 36-mer

SEQ ID NO:31

4. SpeB reverse primer; adds stop site (Amber) and maintains BamHI site:

5' GAATTCGGATCCGCTAGCCTACAACAG 3' 27-mer